

TUNA & Its Open Source Mirrors Infrastructure

Shengqi CHEN 陈晟祺

shengqi.chen@tuna.tsinghua.edu.cn

Open Source @ SIEMENS 2023

2023/11/29 @ Wuxi

Outline

- About TUNA
- About Tsinghua Open Source Mirror
- Current Infrastructure of Tsinghua Open Source Mirror

About TUNA

TUNA is...

- Tsinghua **U**niversity **TUNA A**ssociation
- Formal name: 清华大学学生网络与开源软件协会
- Influential open-source community in mainland China
- Members:
 - In-school: ~500 registered, ~150 in IM group, ~50 active
 - Public: ~1600 in IM group (on Telegram), more on IRC, matrix, etc.

TUNA makes...

- Open-source mirrors:
 - TUNA Mirrors @ THU: mirrors.tuna.tsinghua.edu.cn
 - BFSU Mirrors @ BFSU: mirrors.bfsu.edu.cn
 - OpenTUNA @ AWS: opentuna.cn
 - MirrorZ @ CERNET: mirrorz.org / mirrors.cernet.edu.cn
- Other in-campus services:
 - Public DNS: 101.6.6 / 2001:da8::666
 - NTP: ntp.tuna.tsinghua.edu.cn
 - Livestreaming + danmaku: live.tuna.tsinghua.edu.cn

TUNA hosts...

- Periodic event: Tunight (金枪鱼之夜)
- Normally held on Saturday nights
- Talks about everything!
 - Programming languages, DevOps, HPC, embedded systems, cryptographic, filesystem, OS, ML, mathematics, academic writing, etc.
- Recorded & archived on <https://tuna.moe>

活动公告

金枪鱼之夜: Moonbit 语言及平台

时间: 2023-11-02 19:00-21:00 地点: 三教 3303 Finished

从最初仅以加速 Web 应用为设计目标, WebAssembly 经过数年的发展, 现在已经变为一个通用的跨平台字节码格式。然而, 目前缺乏专门为 WASM 优化的编程语言: C++、Rust 等底层语言语言有陡峭的学习曲线, 生成 WASM 需要很长的编译时间, 而 Golang 等高层语言难以输出高效、紧凑的 WASM 代码。WASM 的高性能、安全性、通用性等优势无法被充分发挥出来。Moonbit 是一个针对 WASM 作为编译目标而设计的通用编程语言, 主要设计目标是编译...

金枪鱼之夜: 2023 年秋季学期迎新会 & TUNA 镜像站服务架构演进

时间: 2023-10-14 19:00-21:00 地点: 旧经管报告厅 Finished

2023 年秋季百团大战胜利(?)收官, 本次招新有很多新的小伙伴加入, 院系分布也更加多元, 欢迎你们加入 TUNA! 本周六将举办迎新会, 按照惯例首先会有 TUNA 的简介、互相认识的环节和建设 TUNA 的 n+1 种方式, 在此之外还会由橙橙分享 TUNA 镜像站的历史和架构演进过程。从 2003 年 ftp.tsinghua.edu.cn 的建立以来, 清华大学的开源镜像站已经走过了二十个年头。这二十年中, 镜像站经历了几代维护者的传承和软硬件架构的不断演进, 到现在...

2023 年秋季学期社团联合招新 (百团大战)

时间: 2023-09-27 / 28 中午及傍晚 地点: 紫操西侧 H4 摊位 Finished

春季学期百团大战就要到了, 又是一年两度决定 TUNA 生死存亡的时候了! 请注意, 由于国庆假期, 本次招新日期是第二周三 (9/27) 及周四 (9/28), 时间依旧是中午及傍晚。活动地点在紫操西侧 H4 摊位, 欢迎一起来吃和玩! 也请各位留意天气情况, 避免被放鸽子。

金枪鱼之夜: 数学/物理演讲稿所见即所得的轻量级编辑方式

时间: 2023-05-27 19:00-21:00 地点: 三教 3104 Finished

资源链接: [📄](#) [📺](#)

在数学和物理领域, 思考往往需要在至少二维的视角进行, 这是因为大量的数学公式都以二维表达, 例如分数, 而物理学更多以画画而非写字来传达思想。遗憾的是, 通常意义的书写方式, 例如用 LaTeX 进行编辑, 是一维的。对于数理工作者, 尤其在准备演讲稿的时候, 所见即所得的编辑模式十分关键。本次 Tunight 将介绍一些比 MS Office 更轻量级的工具, 包括使用 IPE 和 Inkscape 制作报告、结合 LibreOffice 和 LaTeX, 以及在 LyX 中边思考边推导...

TUNA plays...

- Software / Hardware Freedom Day
- Invited Celebrity Talks
 - Author of 《鸟哥的 Linux 私房菜》
 - Richard Stallman
- Open-source conferences
 - GNOME.Asia Summit
 - LC³ Conference
- Distro release parties
- Other LUGs / communities

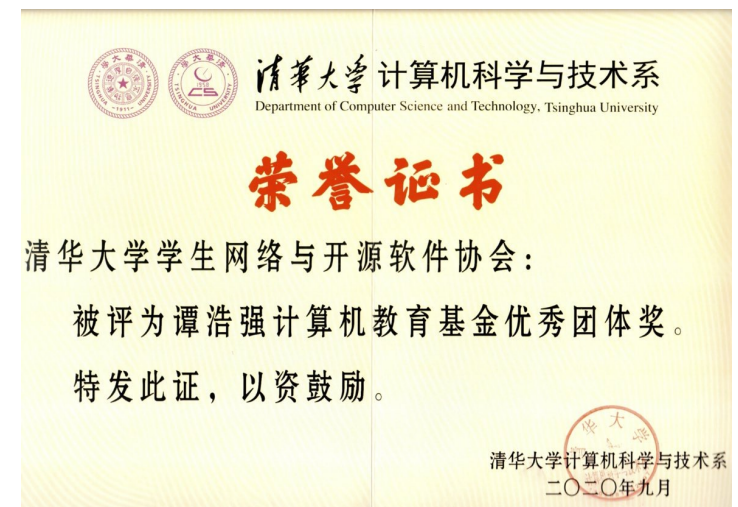


TUNA contributes...

- Mirror solution:
 - tunasync: mirror job management tool in Go
 - mirror-web: frontend for mirror site built with Jekyll
 - mirrorz: info aggregation and auto-redirection for all mirror sites
- tunaccount: shared user database simpler than LDAP
- ThuThesis: LaTeX thesis template for Tsinghua University
- danmaQ: danmaku C/S solution for livestreaming
- fishroom: message forwarder for multiple IM protocols
- QSerial: cross-platform serial port utility
- ...and more, all on <https://github.com/tuna/>

TUNA wins...

- Open Source Promotion Plan (OSPP), 2021
 - Best Quality Award (ZHANG Qifu, XU Pengcheng)
 - Outstanding Contribution Award (WU Yiyang, XU Benda)
- The **First** TAN Haoqiang Computer Education Foundation Award for Outstanding Groups

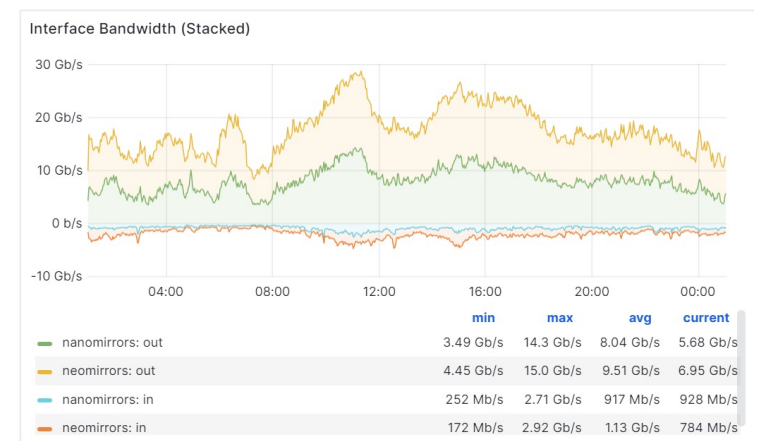


About Tsinghua Open Source Mirror

- {http,https,rsync}://mirrors{4,6}.tuna.tsinghua.edu.cn
- > **180** upstreams: distro / software / font / source code / etc.
- Two load-balancing backend servers
 - Donated by Megvii Technology Ltd., and Huawei Technologies Co., Ltd.
 - Each equipped with ~**100TB all-NVMe** storage
- **40 Gbps** IPv6 bandwidth + **26 Gbps** IPv4 bandwidth
- Managed by a small core team in TUNA: ~5 people
 - @Harry-Chen, @shankerwangmiao, @happyaron, @jiegec

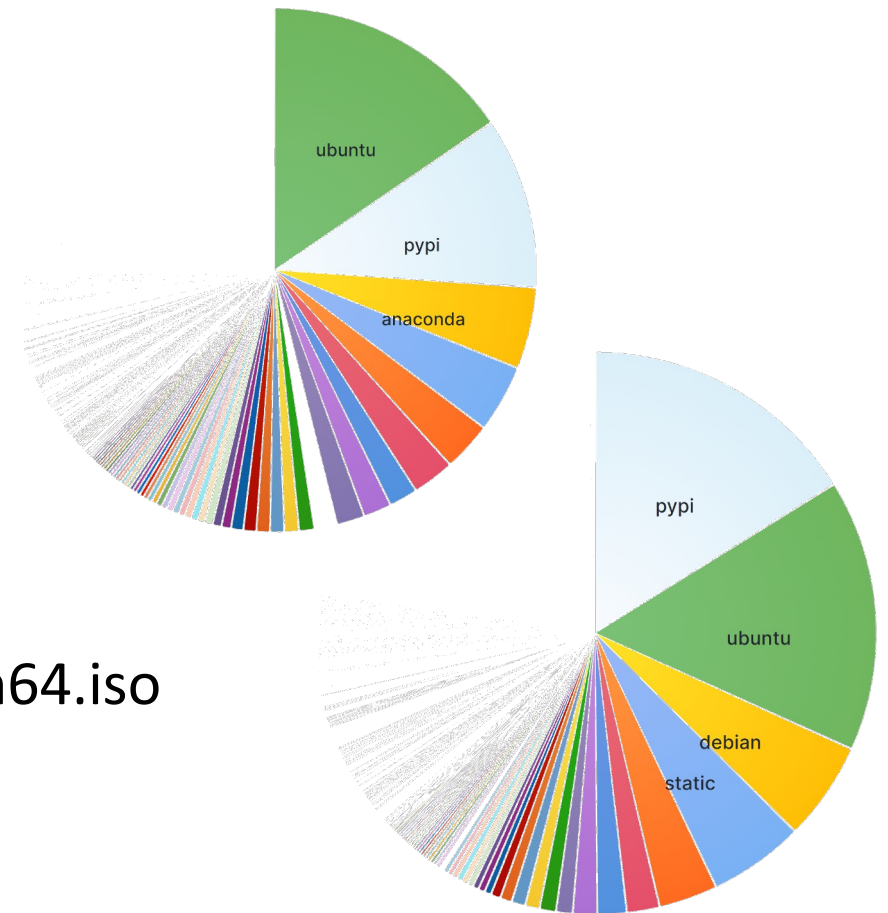
TUNA Mirrors: Statistics

- All data is from 2023/11/23
- Bandwidth usage:
 - IPv4 (w/ xlat): 14.5 Gbps in average, 23 Gbps max
 - IPv6: 2.8 Gbps in average, 5 Gbps max
 - Total: 17.5 Gbps in average, 28.9 Gbps max
- Service distribution:
 - 50% HTTPS, 20% HTTP, 30% rsync in average
 - rsync < HTTP{,S} at night, vice versa



TUNA Mirrors: More statistics

- Top-3 repos by bytes:
 - Ubuntu: 24 Mreq, 31.3 TB
 - Anaconda: 2.8 Mreq, 10.04 TB
 - CentOS: 5.11 Mreq, 8.29 TB
- Top-3 repos by request:
 - PyPI: 25 Mreq, 1.54 TB
 - Ubuntu: 24 Mreq, 31.3 TB
 - Debian: 9Mreq, 672 GB
- Top-1 file: ubuntu-23.10-live-server-arm64.iso
 - 1.28 Kreq, 2.96 TB

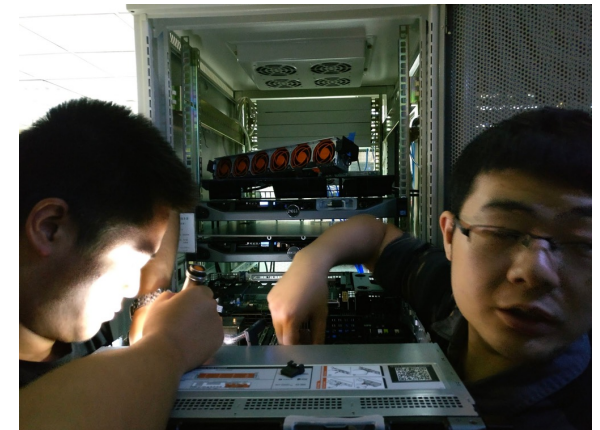


Brief history of TUNA Mirrors

- Pre-history: 2003-2014
 - ftp{2,3,4}.tsinghua.edu.cn: numerous maintainers & tech stacks
 - not so much stability guarantee due to equipment & human factors
- Rebuild: 2014-2016
 - led by previous president bigeagle (大鷹)
 - develop tunasync, the all-in-one job manager for mirrors
 - experienced trials and errors
 - filesystem: btrfs performance issue
 - network: 10G IPv6 + 100M IPv4
 - server: borrowed from university faculty that needed to be returned soon

Opportunity in 2016

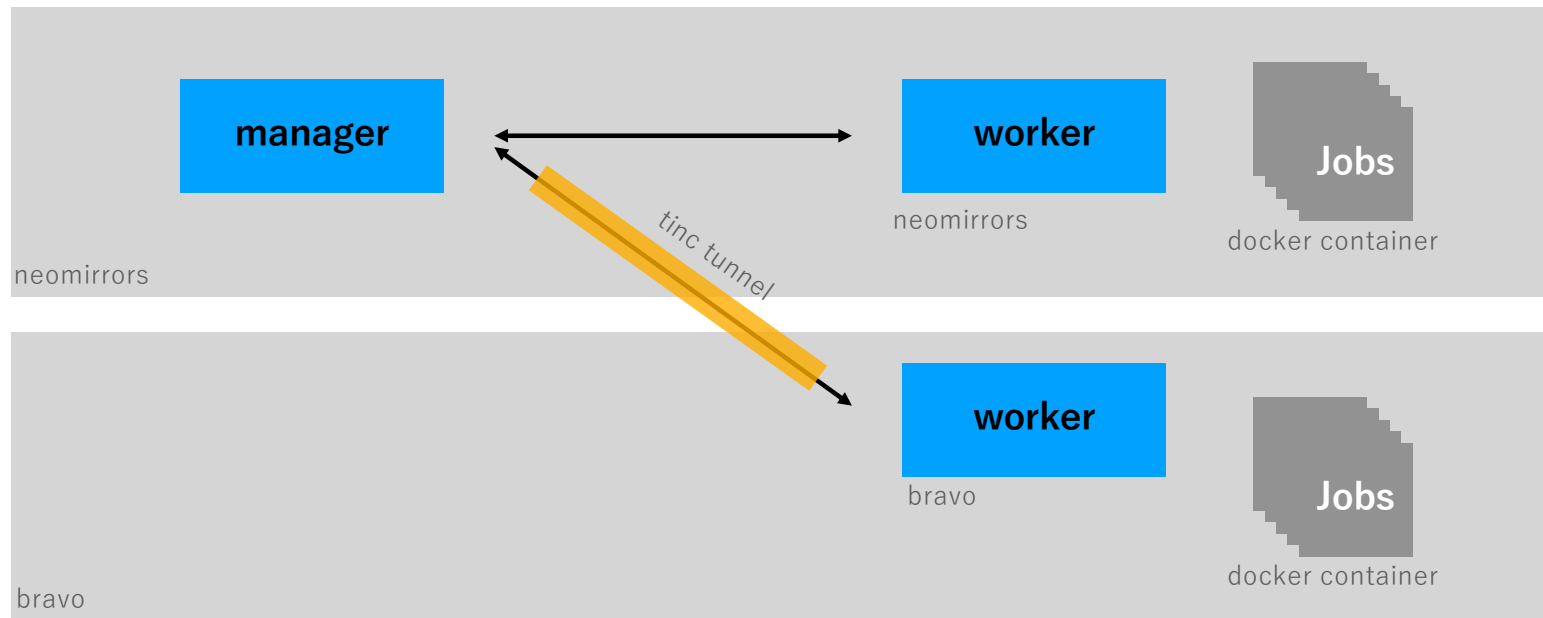
- A former maintainer of ftp2 (youngcow) now works in ITC of Tsinghua University.
- Youngcow saved TUNA mirrors:
 - Run mirrors as official school IT service in well-maintained server room
 - Provided two R720xd servers, each with 12*3T HDD, 1000 Mbps dual stack network
- And Linux Foundation helped too:
 - Donated 4 Dell R620 / R720 servers
 - Re-assembled to 2 servers that runs NS, monitor, git



Software architecture in 2016

- Storage:
 - nanomirrors: 12 * 3T SAS HDD, with hard-RAID6 (~36T usable)
 - neomirrors: 12 * 4T SAS HDD, with **ZFS RAIDZ2** (~40T usable)
- Struggled with IOPS issue, especially caused by rsync
 - developed rsync-huai that caches metadata in memory to alleviate
- Mirror management: tunasync was now rewritten in Go
 - and works with latest technology Docker!
- Network: provide unified access to users
 - use keepalived for hot-backup and switch-over

tunasync with docker



Dual-server Hot-backup mechanism (1)

Based on VRRP

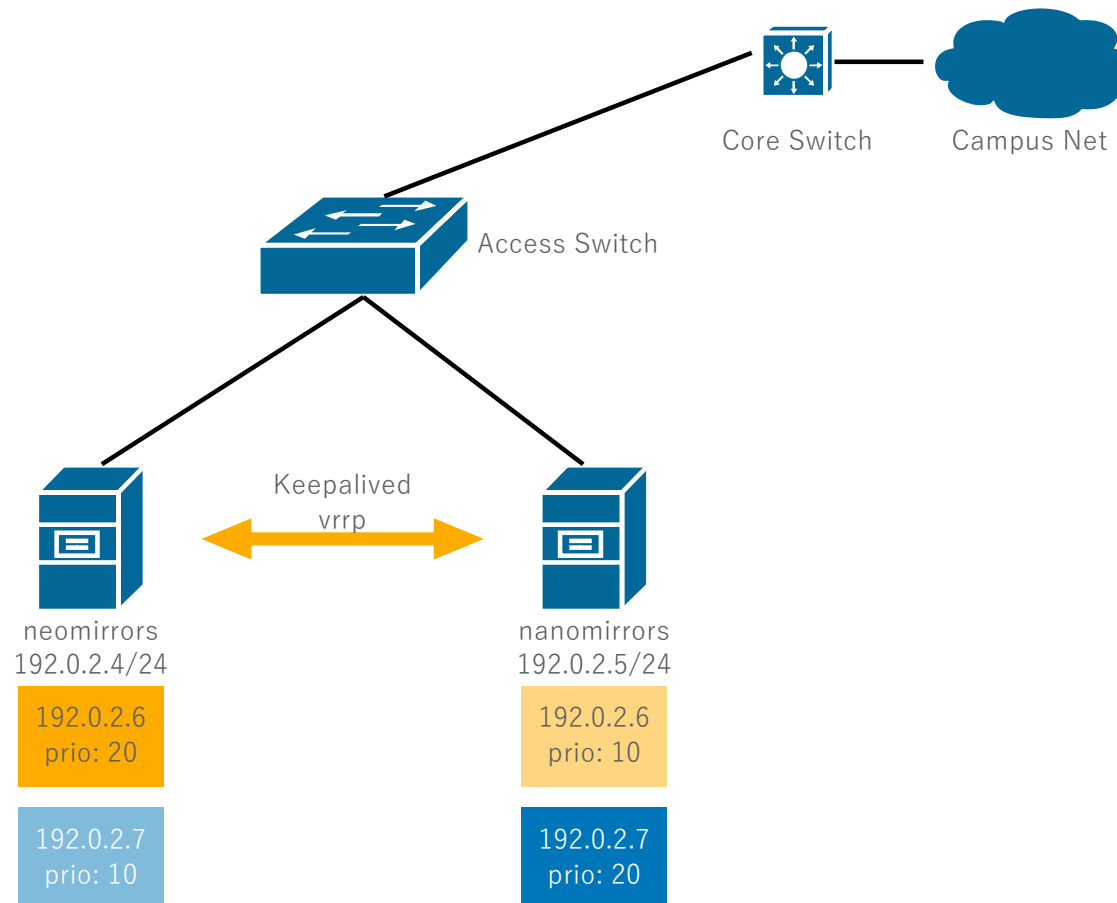
DNS:

mirrors.tuna.tsinghua.edu.cn

Random:

192.0.2.6

192.0.2.7



Dual-server Hot-backup mechanism (1)

Based on VRRP

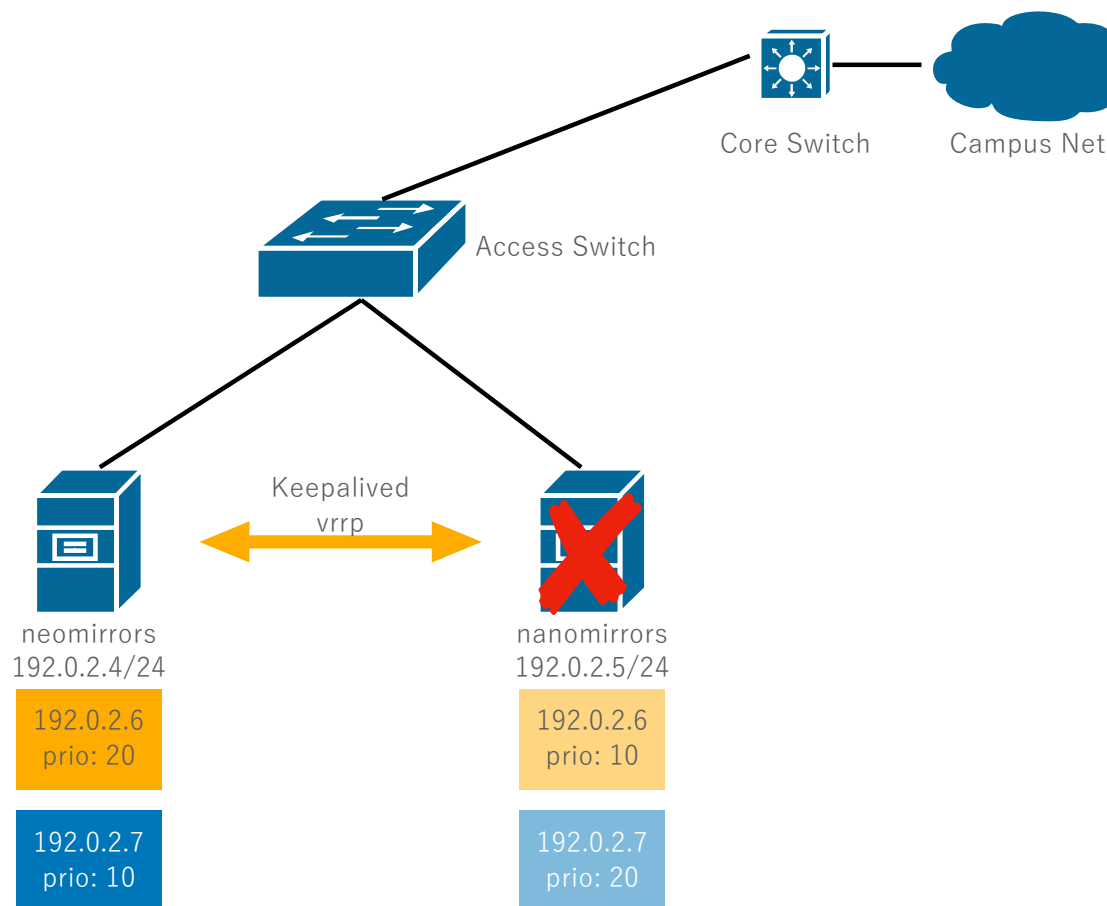
DNS:

mirrors.tuna.tsinghua.edu.cn

Random:

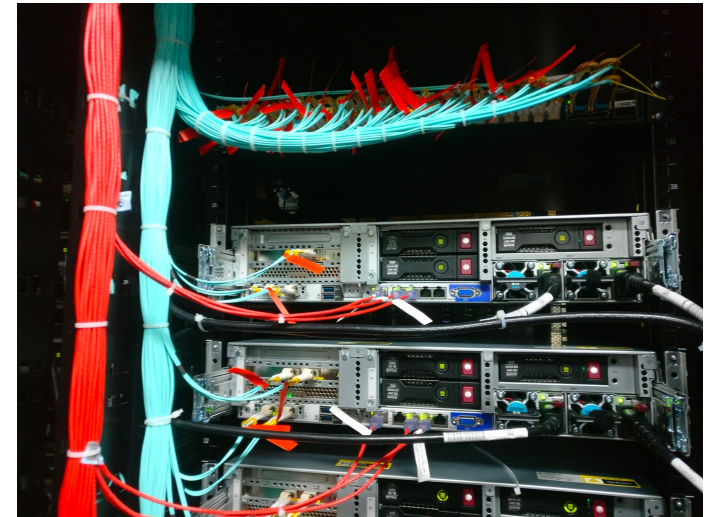
192.0.2.6

192.0.2.7



2018: To the Cloud!

- Two bare-metal HPE server in new cloud platform
- Each with 11×10TB SATA HDD and 2×SATA SSD
 - neomirrors: 11×10TB RAID6 with 2 SSD as cache (~90TB usable)
 - nanomirrors: 2×6×10TB RAIDZ with 2 SSD as L2ARC (~100TB usable)
- Each with 10Gbps campus network
 - hard to saturate still due to IOPS issue

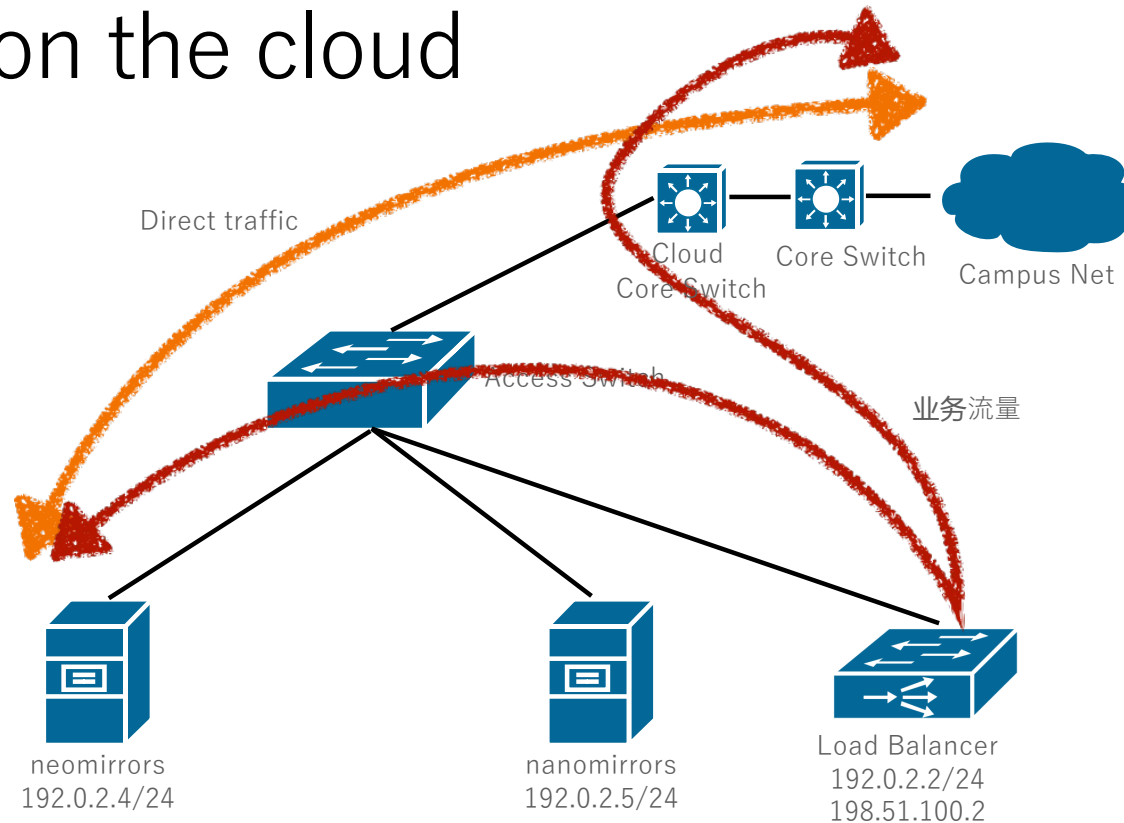


Load-balancer on the cloud

DNS:

mirrors.tuna.tsinghua.edu.cn

198.51.100.2



```
iptables:
match mac src = lb, connmark 100
match conn mark 100, mark 100

ip rule:
match mark 100 lookup table 100

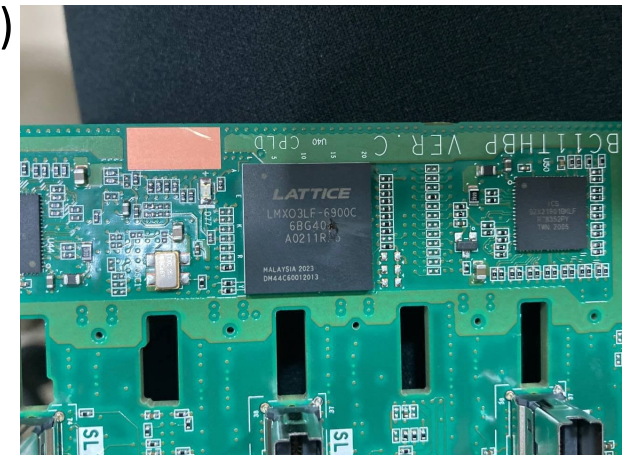
route 0/0 via 192.0.2.2 table 100
route 0/0 via 192.0.2.1 table main
```

Load-balancer is NO silver bullet.

- The LB does a layer-4 translation that:
 - becomes a critical path between users and servers
 - has FPGA that accelerates forwarding.
- BUG#1: NAT entry in FPGA vanished 30s after connection established.
 - User experienced disruption when downloading large files.
 - Workaround: disable hardwired logic in FPGA, use CPU to forward
- BUG#2: unbearable package loss of inbound TCP ACK
 - Download speed reduces drastically after certain time.
 - Meme:“请你们体面地关闭镜像站”
 - Workaround: **NO**

2020-2022: Back on ground again

- In Aug 2020 and May 2021, Megvii and Huawei donated two all-NVMe servers to TUNA:
 - One is traditional X86-64, the other is aarch64 (Kunpeng 920)
 - Each with 20×6.4T NVMe storage and 2×10Gbps network
 - All in ZFS finally: 4×5×6.4T RAIDZ (~96T usable)
- Fun fact: We burnt the NVMe backplane when modifying servers due to incompatible cable
- IOPS is now finally no longer a problem.
- And we do not want LBs any more.



Dual-server hot-backup mechanism (2)

Based on ECMP and BFD

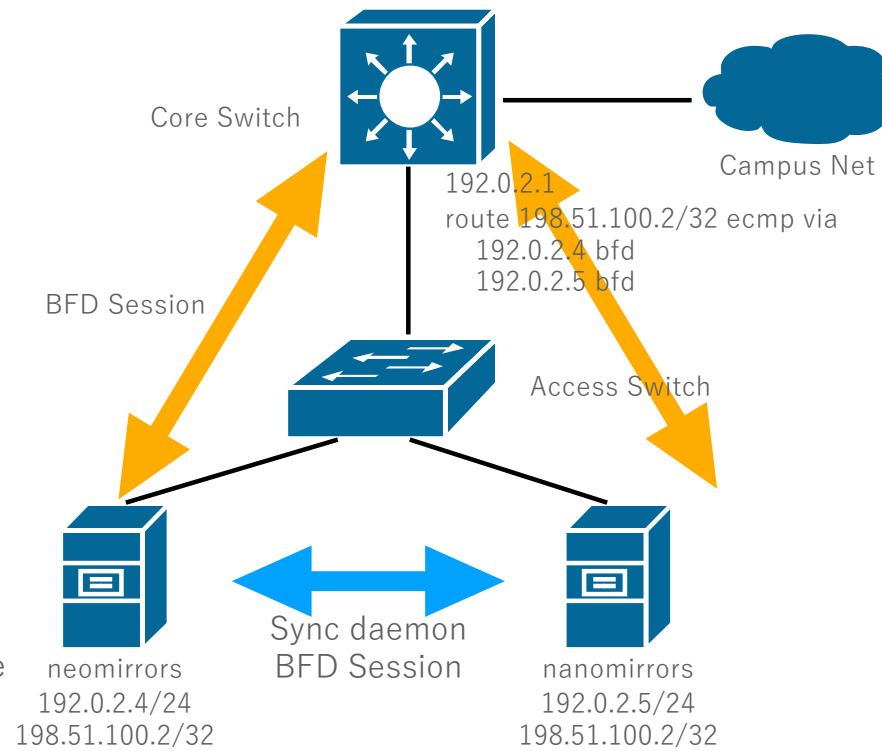
DNS:

mirrors.tuna.tsinghua.edu.cn

198.51.100.2

ipvs:

vs 198.51.100.2 source hash, route
rs1: 192.0.2.4
rs2: 192.0.2.5



ipvs:

vs 198.51.100.2 source hash, route
rs1: 192.0.2.4
rs2: 192.0.2.5

Compare VRRP and ECMP+BFD

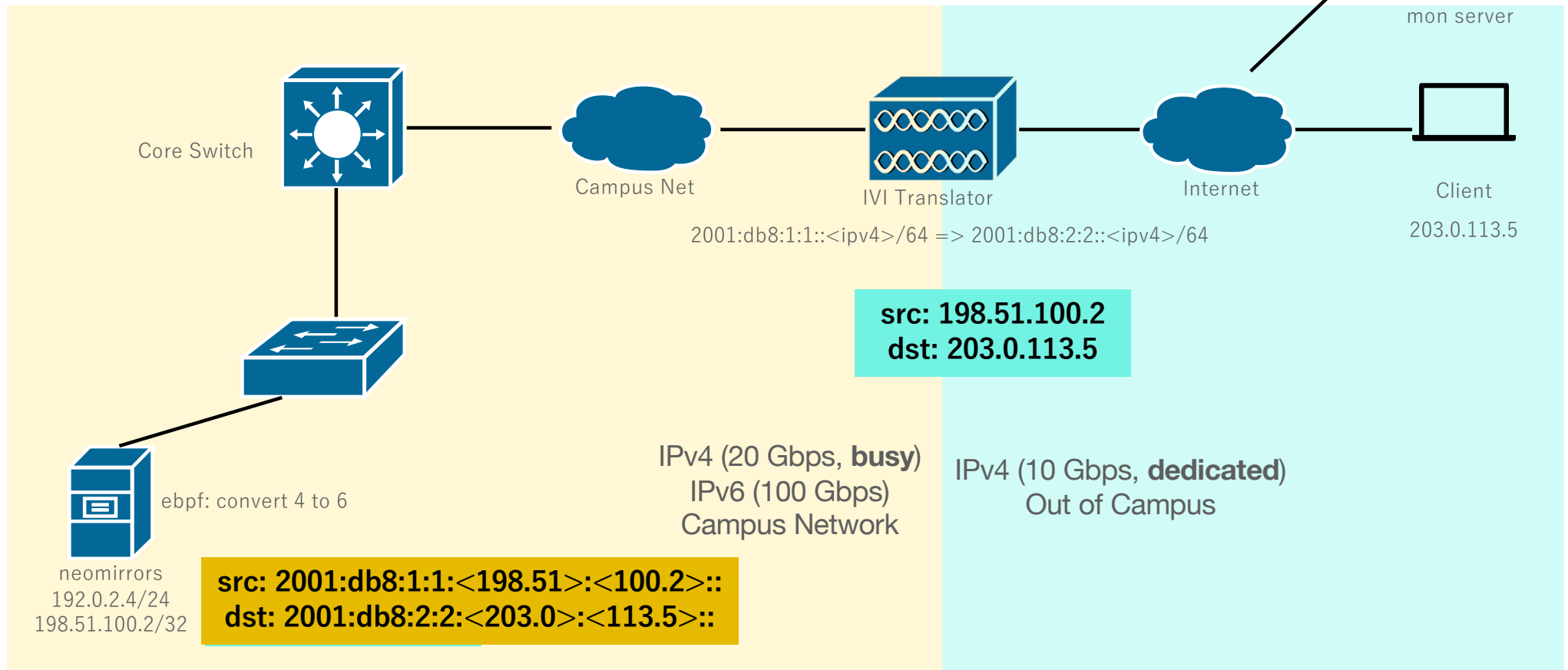
	VRRP	ECMP + BFD
Load-balance	DNS record (layer 7)	ECMP (layer 3)
Auto switch-over on failure	Yes	Yes
Traffic overhead	None	None
Smooth manual switch	No	Yes
Consistency for same user	Partial (by external methods like DNS cache)	Yes (with IPVS)

2023: leverage IPv6 resource

- Mirror traffic saturated IPv4 campus network.
- The total bandwidth of Tsinghua campus network in 2022:
 - IPv4: 20Gbps, contended
 - IPv6: 100Gbps, vacant
- Could we transfer IPv4 traffic through IPv6 network?
 - Yes! With the help of IIVI translation technology (RFC 6144, RFC 6219)

xlat464

Single-side IVI translation



Conclusion

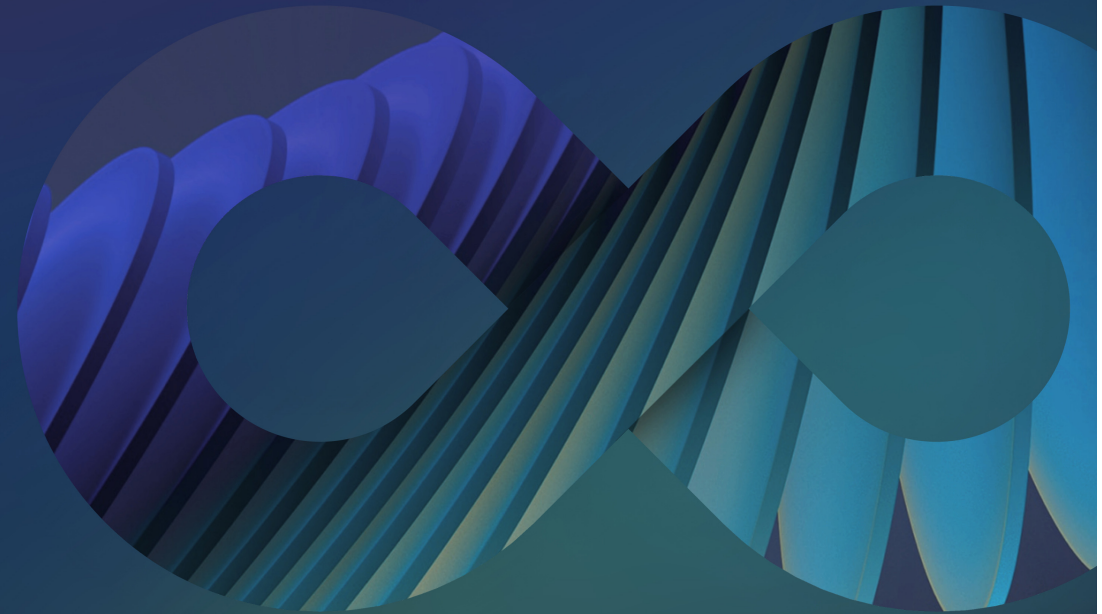
- TUNA mirrors has gone through 20 years, now with:
 - 102.4 TB high-speed fully-redundant NVMe storage
 - 32 TB flash storage (SATA SSDs)
 - 40 Gbps IPv6 bandwidth
 - 20 Gbps xlat-ed IPv4 bandwidth + 6 Gbps IPv4 bandwidth
- Thanks to all sponsors, supporters, and users!

Thanks

Q & A

Open Source Infrastructure

Open Source @ 西门子
Shengqi Chen & Felix Moessbauer



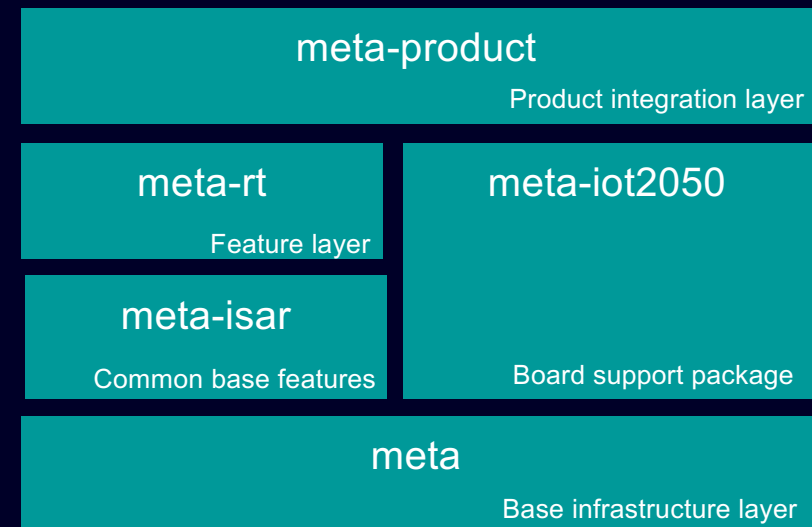
A Peek at Meta Builders

KAS

- Tool to run build in isolated and deterministic build environment
- Manage fetching and patching of build recipes ("layers")
- Executes ISAR build

ISAR

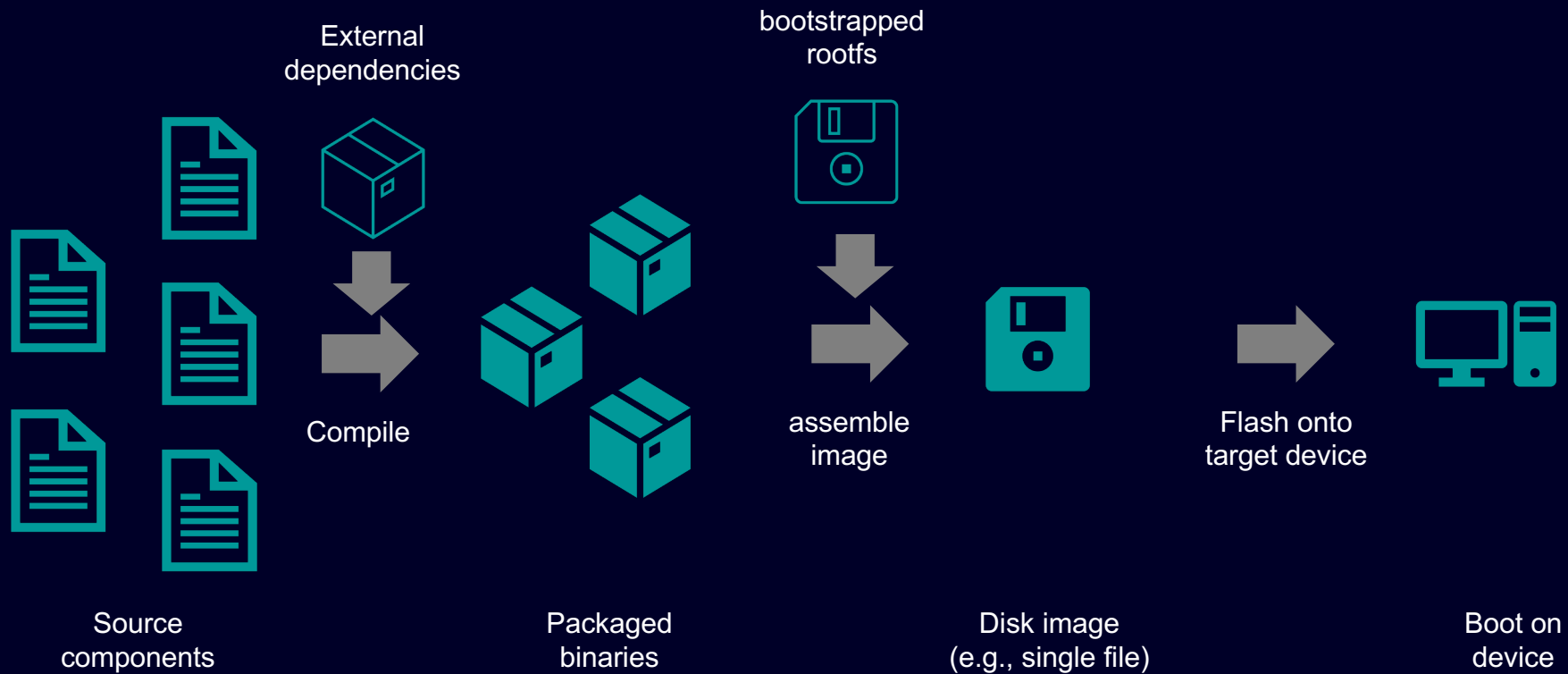
- Like Yocto / OpenEmbedded
- Provides recipes to describe image building



Example layers for an embedded image build. All boxes are individual git repositories.

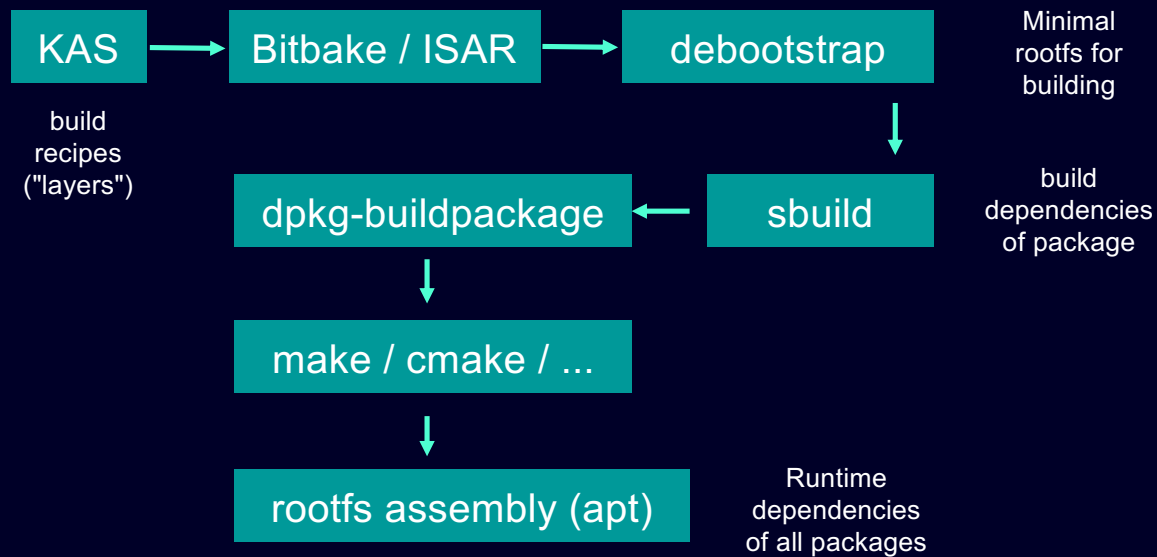
Building bootable images for embedded devices @Siemens

Example of a simplified build pipeline



Building Images for Embedded Systems

A massively simplified example of an embedded image build using ISAR + KAS



Build systems are layered

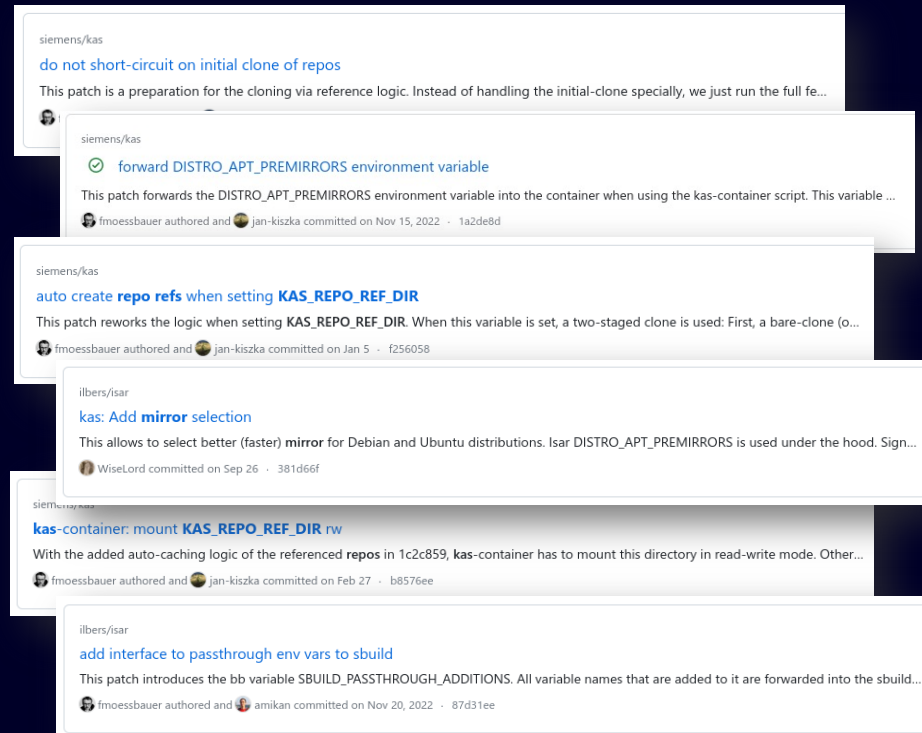
For embedded image building a whole chain of build tools is used to build components

Tools

- Meta builders (KAS)
- Bitbake / Yocto / ISAR
- Debian: dpkg-buildpackage / sbuild
- C / C++: autotools, cmake, meson
- NPM / Ruby

Improve Local and CI/CD Build Performance

Key Enablers



Selection of commits to improve or add mirror support in various tools.

Improving build performance

Extensive Caching

- Cache all downloads
- Cache all layers (git repos)
- Cache all build artifacts (SSTATE_CACHE)

Using Chinese mirrors

- Use local mirrors whenever possible
- Fast and reliable access needed in CI
- Document available mirrors and how to use them

Documentation is Key

Make knowledge about mirrors available to your developers

Mirror injection depends on tool / programming language

- Debian / APT
- Docker
- Go
- Linux Kernel
- NPM
- Ruby
- Rust / Cargo

If you want to apply these mirrors, you need to update your `/etc/apt/sources.list` file.
For example, for Ubuntu 22.04 this could look as follows:

```
1 deb http://mirrors.tuna.tsinghua.edu.cn/ubuntu jammy main
2 deb http://mirrors.tuna.tsinghua.edu.cn/ubuntu jammy-updates main
3 deb http://mirrors.tuna.tsinghua.edu.cn/ubuntu jammy-security main
4 deb http://mirrors.tuna.tsinghua.edu.cn/ubuntu jammy-updates-security main
```

To use these mirrors, you need to update your `/etc/docker/registry/mirrors` and Ubuntu):

```
{
  "registry-mirrors": [
    "https://docker.mirrors.ustc.edu.cn"
  ]
}
```

The following configuration is required to use goproxy.cn:

```
1 export GO111MODULE="on"
2 export GOPROXY="https://goproxy.cn,direct"
```

To activate the mirror for ruby gems:

```
1 gem sources --add https://mirrors.tuna.tsinghua.edu.cn/rubygems/ --remove=ht
```

To activate the mirror for Bundler:

```
1 bundle config mirror.https://rubygems.org https://mirrors.tuna.tsinghua.edu.
```

The activation of a Cargo mirror is done by modifying (or creating) a `config.toml` file with the following content:

```
1 [source.crates-io]
2 registry = "https://github.com/rust-lang/crates.io-index"
3 replace-with = 'tuna'
4
5 [source.tuna]
6 registry = "https://mirrors.tuna.tsinghua.edu.cn/crates-registry-index"
```

The installation of Rust via a mirror is done in the following way:

```
1 RUSTUP_DIST_SERVER=https://mirrors.tuna.tsinghua.edu.cn/rustup rustup init
```

For further details, check the [source replacement specification](#) of Cargo.

Not publicly available anymore

Screenshots from Siemens developer portal describing how to configure the mirrors for various tools and environments.

Q&A

Contact:

Felix Moessbauer

Siemens Ltd. China

T CED INW-CN

E-mail felix.moessbauer@siemens.com

Shengqi Chen

TUNA Project - Tsinghua University Beijing

E-mail shengqi.chen@tuna.tsinghua.edu.cn

Open Source Resources:

- <https://mirrors.tuna.tsinghua.edu.cn/>
- <https://github.com/ilbers/isar/>
- <https://github.com/siemens/kas/>